(19)日本国特許庁（ＪＰ）　(12) **公開特許公報**（Ａ）　(11)特許出願公開番号

特開平9－288547

(43)公開日　平成９年(1997)11月４日

(51)Int.Cl.[6]	識別記号	庁内整理番号	ＦＩ		技術表示箇所
Ｇ０６Ｆ　3/06	５４０		Ｇ０６Ｆ　3/06	５４０	
	３０１			３０１Ｊ	

審査請求　未請求　請求項の数４　ＦＤ　（全 15 頁）

(21)出願番号　特願平8－124088

(22)出願日　平成８年(1996)４月22日

(71)出願人　000153454
株式会社日立インフォメーションテクノロジー
神奈川県秦野市堀山下１番地

(72)発明者　吉田　徹
神奈川県秦野市堀山下１番地　株式会社日立コンピュータエレクトロニクス内

(72)発明者　青木　健
神奈川県秦野市堀山下１番地　株式会社日立コンピュータエレクトロニクス内

(72)発明者　舘下　昌司
神奈川県秦野市堀山下１番地　株式会社日立コンピュータエレクトロニクス内

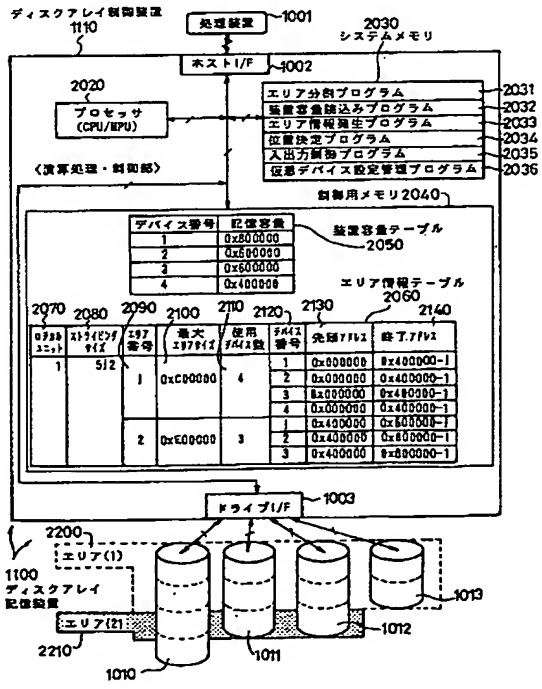
(74)代理人　弁理士　梶山　佶是　（外１名）

(54)【発明の名称】　アレイ型記憶装置

(57)【要約】　（修正有）

【課題】複数の記憶デバイスの１つに冗長データを記憶し、複数のデバイスに共通のアドレス空間を割当てて単一の記憶装置として管理するアレイ型記憶装置において、搭載する記憶デバイスの記憶容量の制限をなくすこと。

【解決手段】複数の記憶デバイスのそれぞれの最大記憶容量を読出す手段と、記憶デバイスの記憶容量のうち最も小さい第１の最大記憶容量に合わせて第１のアドレス空間を割当てアレイを形成する第１の記憶エリアを設定し、かつ、最も小さい第１の最大記憶容量の記憶デバイスを除いた複数の記憶デバイスにおいてそれぞれの最大記憶容量から最も小さい第１の最大記憶容量分を差引いた残りのうち最も小さい第２の最大記憶容量に合わせて第２のアドレス空間を割当てアレイを形成する第２の記憶エリアを設定するエリア設定手段と、上位処理装置からのアクセスアドレスに基づいてアレイを形成する第１または第２の記憶エリアに従って記憶デバイスの物理的なアドレスを算出するアドレス算出手段とを備える。



【特許請求の範囲】
【請求項１】複数の記憶デバイスの１つに冗長データを記憶し、複数のデバイスに共通のアドレス空間を割当てて単一の記憶装置として管理するアレイ型記憶装置において、前記複数の記憶デバイスのそれぞれの最大記憶容量を読出す手段と、前記記憶デバイスの記憶容量のうち最も小さい第１の最大記憶容量に合わせて第１の前記アドレス空間を割当てアレイを形成する第１の記憶エリアを設定し、かつ、前記最も小さい第１の最大記憶容量の前記記憶デバイスを除いた複数の前記記憶デバイスにおいてそれぞれの最大記憶容量から前記最も小さい第１の最大記憶容量分を差引いた残りのうち最も小さい第２の最大記憶容量に合わせて第２の前記アドレス空間を割当てアレイを形成する第２の記憶エリアを設定するエリア設定手段と、上位処理装置からのアクセスアドレスに基づいて前記アレイを形成する前記第１または第２の記憶エリアに従って前記記憶デバイスの物理的なアドレスを算出するアドレス算出手段とを備えるアレイ型記憶装置。
【請求項２】前記複数の記憶デバイスのうちの２以上のデバイスを１つの仮想記憶デバイスとして処理する仮想記憶デバイス設定手段を備え、前記エリア設定手段は、この仮想記憶デバイス設定手段により仮想記憶デバイスとされた前記２以上の記憶デバイス以外のデバイスも仮想記憶デバイスとして各前記仮想記憶デバイスについて前記第１および第２のエリア設定をするものであり、前記アドレス算出手段は、前記仮想記憶デバイスに基づいて前記記憶デバイスの物理的なアドレスを算出するものである請求項１記載のアレイ型記憶装置。
【請求項３】複数の記憶デバイスと、上位処理装置のデータ書込み要求に対して前記上位処理装置の書込みデータに冗長データを付加し、前記書込みデータと冗長データとを前記記憶デバイスに分散して書込み、前記上位処理装置からのデータ読出し要求に対して必要なデータを前記記憶デバイスから読出して前記上位処理装置に転送し、かつ、障害などにより前記記憶デバイス上のデータに直接アクセスできない場合に前記冗長データを用いてアクセスできないデータを復元する制御装置を有するアレイ型記憶装置において、前記記憶デバイスの単体容量を前記制御装置に読出す読出手段と、前記読出手段で読出した前記記憶デバイスの単体容量に基づいて前記複数の記憶デバイスに所定のアドレス空間を割当て複数のアレイとして管理する管理手段と、前記上位処理装置からアクセスされたときに、前記上位処理装置からのアクセスアドレスに基づいて前記複数のアレイの１つを形成する前記記憶デバイスの物理的なアドレスを算出するアドレス算出手段とを備えるアレイ型記憶装置。
【請求項４】前記管理手段は、前記記憶デバイスに対してアレイとして管理する際に、２台以上の記憶デバイスを仮想的な１台の記憶デバイスとして管理するものであり、前記アドレス算出手段は、前記上位処理装置からアクセスされたときに、前記上位処理装置からのアクセスアドレスに基づいて前記仮想的な記憶デバイスの論理的なアドレスを算出する論理アドレス算出手段と前記論理的アドレス算出手段で算出された論理的なアドレスから前記記憶デバイスの物理的なアドレスを算出する物理的アドレス算出手段とを有するアレイ型記憶装置。
【発明の詳細な説明】
【０００１】
【発明の属する技術分野】この発明は、アレイ型記憶装置において、ファイルなどのまとまりのあるデータを複数の磁気ディスク記憶装置等の外部記憶装置（以下記憶デバイスという）に分散して冗長データを付加して記憶するアレイ型の記憶装置に関し、詳しくは、容量の異なる記憶デバイスを複数使用した場合においても、アレイ形式で効率よく管理することができるアレイ型の記憶装置に関する。
【０００２】
【従来の技術】従来のアレイ形式の記憶装置として磁気ディスク記憶装置を多数配列してこれらを一体的な単一の記憶領域として扱うディスクアレイ・システムが一般的に知られている。ディスクアレイ・システムは、計算機システムにおいて、多数の磁気ディスク装置を並列に入出力動作させることにより、処理装置と記憶デバイスとの間で高速なデータ転送を行い、高い信頼性の実現をするシステムである。
【０００３】ディスクアレイ・システムの構成の例としては、Ｐａｔｅｒｓｏｎ等の論文（D. Patterson, G. Gibson, R. Katz,” A Case for Redundant Arrays of InexpensiveDisks(RAID)”, ACM SIGMOD conference proceedings, 1988, pp.109-116）に記載がある。この論文では、ＲＡＩＤ（レイド，Ｒｅｄｕｎｄａｎｔ Ａｒｒａｙ ｏｆ Ｉｎｅｘｐｅｎｓｉｖｅ Ｄｉｓｋｓ）と呼ばれるディスクアレイ形態を提示している。ＲＡＩＤは、通常の入出力データを複数の記憶デバイスに分散記憶すると共に、冗長データを記憶する。この冗長データの設置は、前記データに欠落が生じた場合に、データ復元処理を行うために用いられ、これにより、欠落したデータを復元することを可能にする。
【０００４】
【発明が解決しようとする課題】図７は、アレイ型記憶装置の１つとして一般的に知られているディスクアレイ記憶装置を用いるディスクアレイ記憶システムの機能ブロック図である。図７は、アレイ型の記憶システム（ディスクアレイ記憶システム）を示していて、これは、入出力要求を発行する、プロセッサを有する処理装置１００１（いわゆるホストコンピュータ）と各種のデータを記憶するディスクアレイ記憶装置１０３０とからなる。ディスクアレイ記憶装置１０３０は、ディスクアレイ制御装置１０００とデータを記憶する磁気ディスク記憶装

置（以下ディスク記憶デバイスという。）１０１０～１０１３等から構成されている。ディスクアレイ制御装置１０００は、処理装置１００１と物理的に接続されるホストインタフェース（以下インタフェースはＩ／Ｆと略す。）１００２、制御用メモリ１００４、ディスク記憶デバイス１０１０～１０１３との間で物理的な接続を行うドライブＩ／Ｆ１００３、そして冗長データの生成、データの読出／書込制御等の各種の制御を行なうための演算処理・制御部（ＣＰＵ＋メモリ）１００５とにより構成される。

【０００５】前記の構成によりディスクアレイ制御装置１０００は、処理装置１００１からのアクセスに対して、ディスク記憶デバイス装置１０１０～１０１３との入出力処理、冗長データの生成、冗長データを用いたデータの復元処理などを行なう。また、ロジカルユニット１０２０は、ディスク記憶デバイス１０１０～１０１３を処理装置１００１からみて論理的に１台の記憶装置デバイスとして認識される記憶領域を示すものであって、１０２１は、ディスク記憶デバイス１０１０～１０１３において処理装置１００１から１台の記憶装置として認識されない未使用領域である。

【０００６】このようなディスクアレイ記憶装置にあっては、記憶されるデータの冗長性を維持し、かつ搭載される記憶デバイスの記憶容量を最大限に利用するために、搭載する記憶デバイスの各々の記憶容量を同一にする必要がある。そのため、図７に示すように搭載される記憶デバイスの中に一つでも記憶容量の少ない記憶装置があった場合には、記憶容量が最小の記憶デバイスを基準にしてアレイ型記憶装置１０３０が構成される。それが論理的に構築する１つのディスク記憶デバイス（ロジカルユニット）１０２０である。すなわち、最小の記憶容量を持つディスク記憶デバイスより多い記憶容量を持つディスク記憶デバイスは、最小記憶容量との差分の記憶容量が未使用（未使用領域１０２１）となり、この記憶領域（記憶単体）が無駄になる問題がある。

【０００７】また、アレイ型記憶装置において、記憶容量増加の目的で増設されるディスク記憶デバイスも、増設したディスク記憶デバイスを最大限に利用するためには、増設以前に搭載されているディスク記憶デバイスの記憶容量を考慮して同一な記憶容量のものを用意することが必要になる。この発明の目的は、このような従来技術の問題点を解決するものであって、データを複数の記憶デバイスに分散して記憶するアレイ型記憶装置において、記憶デバイスの記憶容量を考慮しなくても済み、搭載する記憶デバイスの記憶容量の制限をなくすことができるアレイ型記憶装置を提供することにある。

【０００８】

【課題を解決するための手段】このような目的を達成するためのこの発明のアレイ型記憶装置の特徴は、複数の記憶デバイスの１つに冗長データを記憶し、複数のデバイスに共通のアドレス空間を割当てて単一の記憶装置として管理するアレイ型記憶装置において、複数の記憶デバイスのそれぞれの最大記憶容量を読出す手段と、記憶デバイスの記憶容量のうち最も小さい第１の最大記憶容量に合わせて第１のアドレス空間を割当てアレイを形成する第１の記憶エリアを設定し、かつ、最も小さい第１の最大記憶容量の記憶デバイスを除いた複数の記憶デバイスにおいてそれぞれの最大記憶容量から最も小さい第１の最大記憶容量分を差引いた残りのうち最も小さい第２の最大記憶容量に合わせて第２のアドレス空間を割当てアレイを形成する第２の記憶エリアを設定するエリア設定手段と、上位処理装置からのアクセスアドレスに基づいてアレイを形成する第１または第２の記憶エリアに従って記憶デバイスの物理的なアドレスを算出するアドレス算出手段とを備えるものである。

【０００９】

【発明の実施の形態】前記の構成のように、この発明にあっては、複数の記憶デバイスの単体容量が異なる場合にあっても、上位処理装置から１台の記憶装置として認識されるエリア（記憶単位）を持つ、異なったアレイ複数を形成する。さらに、他の発明として、アレイを形成する際には、２台以上の記憶デバイスを仮想的な１台の記憶デバイスと見なして、上位処理装置からのアクセスされた際に、上位処理装置からのアクセスアドレスから、仮想的な記憶デバイスの論理的なアドレスを特定し、特定した論理的なアドレスから前記記憶デバイスの物理的なアドレスを特定するようにする。このようにすることで、制御装置の配下に接続される記憶デバイスの単体容量が異なる場合であっても、複数のアレイ領域を持たせることで上位処理装置から１台の記憶装置として認識されるようにすることができる。その結果、異なる記憶容量を持つ記憶デバイスを使用しても、上位処理装置から１台の記憶装置として認識される容量効率の高いアレイ形式の記憶装置を実現することができる。なお、アレイ型記憶装置では、ストライピングサイズ（後述）で分散記憶をするために、冗長データを記憶するデバイスは、複数の記憶デバイスの１つではあるが、この１つの記憶デバイスは、データの書込みの都度決定され、その都度異なるものになる。

【００１０】

【実施例】図１は、この発明のアレイ型記憶装置を適用した一実施例のディクアレイ記憶システムの機能ブロック図である。図１において、１１００は、ディクアレイ記憶装置であって、ディスクアレイ制御装置１１１０とディスク記憶デバイス１０１０～１０１３とにより構成され、ディスクアレイ制御装置１１１０は、演算処理・制御部２０１０と制御メモリ２０４０等を有している。演算処理・制御部２０１０は、プロセッサ（ＣＰＵあるいはＭＰＵ）２０２０と各種処理プログラムを記憶したシステムメモリ２０３０からなる。制御メモリ２０４０

は、図７の制御用メモリ１００４に対応する制御用メモリである。システムメモリ２０３０には、エリア分割プログラム２０３１と装置容量読込みプログラム２０３２、エリア情報発生プログラム２０３３、位置決定プログラム２０３４、入出力制御プログラム２０３５、仮想デバイス設定管理プログラム２０３６等が設けられている。

【００１１】装置容量読込みプログラム２０３２は、ディスクアレイ記憶装置を構築する際に、ドライブＩ／Ｆ１００３に接続された各ディスク記憶デバイス１０１０～１０１３の記憶容量を読込み、制御用メモリ２０４０にディスク記憶デバイス番号別に記憶容量（その記憶デバイスの最大記憶容量）のデータをテーブルの形式（装置容量テーブル２０５０）で格納する処理プログラムである。このプログラムは、その終了時点で次のエリア分割プログラム２０３１をコールする。なお、このとき、ディスク記憶デバイス１０１０～１０１３は、デバイス番号１～４が割当てられ、この番号で管理される。エリア分割プログラム２０３１は、前記のディスク記憶デバイス番号別に作成された記憶容量データを元にして複数のデバイス間においてアドレス空間が共通するエリアを単位として順次エリア分割して複数のエリアを論理的に発生させて番号付けをし、アレイとして発生した各エリアを管理する処理プログラムである。ここでは、ディスク記憶デバイス１０１０～１０１３の記憶容量を元にして２つのエリア(1)２２００とエリア(2)２２１０とを発生させ、各エリアに対応させてディスク記憶デバイス情報をテーブル２０６０の形式で制御用メモリ２０４０に形成する。

【００１２】エリア情報発生プログラム２０３３は、処理装置１００１からのアクセス要求ブロックアドレス（ストライピングサイズ、例えば５１２バイトごとに分割されて領域をアクセスするので、各アドレスはブロックアドレスになる。）を元にして前記のテーブル２０６０を参照してアクセスすべき分割されたエリアを特定するエリア情報（エリア(1)か、エリア(2)か）を発生する処理プログラムである。なお、処理装置１００１は、アクセス要求ブロックアドレスのほかに、記憶情報についてのサイズ情報も送出する。このサイズ情報に応じてエリアが次にまたがる場合には、そのエリア情報も発生する。また、前記のストライピングサイズは、データを各デバイスごとに分散記憶する場合のデータ記憶の単位であって、これはデータ転送の単位として管理される領域、いわゆる１セクタのバイト数に相当するものである。位置決定プログラム２０３４は、処理装置１００１からのアクセス要求ブロックアドレス及び前記エリア情報を元にして、計算式により冗長データを格納するディスク記憶デバイス、データを格納するディスク記憶デバイス及びアクセス要求データのディスク記憶デバイスにおけるブロックアドレス（物理アドレス）を後述する計算式によりそれぞれ算出する処理プログラムである。記憶容量テーブル２０５０は、装置容量読込みプログラム２０３２により形成され、制御用メモリ２０４０上に作成されるデバイスごとの、それぞれの最大記憶容量を記憶するテーブルである。エリア情報テーブル２０６０は、エリア分割プログラム２０３１により形成され、制御メモリ２０４０上に作成されるエリアの詳細情報を記憶するテーブルである。

【００１３】エリア情報テーブル２０６０の欄２０７０は、ロジカルユニット番号の格納場所である。欄２０８０は、ディスクアレイ装置構築の際に指定するストライピングサイズについて格納する場所である。欄２０９０は、分割されたエリアの番号を格納する場所である。欄２１００は、一つのエリアで使用されている論理空間（アドレス空間）の最大エリアサイズ格納場所である。欄２１１０は、一つのエリア内で使用されているディスク記憶デバイスの個数を格納する場所である。欄２１２０はデバイス番号、欄２１３０は先頭物理アドレス、欄２１４０は最終物理アドレスをそれぞれディスク記憶デバイスに対応させて記憶する欄である。２２００は、エリア分割プログラム２０３１により生成されたディスク記憶デバイス１０１０～１０１３を対象として分割して論理的に形成された第１のアレイエリアを構成する１つ目のエリア(1)であり、２２１０は、ディスク記憶デバイス１０１０～１０１３を対象として分割して論理的に形成された第２のアレイエリアを構成する２つ目のエリア(2)である。

【００１４】まず、装置容量読込みプログラム２０３２は、各ディスク記憶デバイス１０１０～１０１３に対してＲｅａｄ Ｃａｐａｃｉｔｙコマンドを発行して、ディスク記憶デバイス毎の記憶容量を読込み、制御用メモリ２０４０内に装置容量テーブル２０５０を設けてデバイス番号に対応させてそれぞれの記憶容量を記憶し、その処理を終了する。エリア分割プログラム２０３１は、装置容量読込みプログラムの終了により、コールされて実行され、次に説明するエリア分割処理とエリア情報テーブル２０６０の生成処理を行う。

【００１５】前記の処理について図２のフローチャートに従って説明すると、まず、エリア分割プログラム２０３１は、始めにロジカルユニット番号をエリアテーブル２０６０内のロジカルユニット番号格納場所である欄２０７０に格納し（ステップ３０１０）、これから構築しようとするディスクアレイ装置のストライピングサイズをストライピングサイズ格納場所２０８０に格納した後に（ステップ３０２０）、ディスク記憶デバイス単体で使用済みの記憶容量格納変数ｚを初期化（ステップ３０３０）し、エリア情報テーブルの内容を作成するために装置情報テーブル２０５０をメモリ２０３０の作業領域にコピーする処理を行う（ステップ３０４０）。

【００１６】次に、最小記憶容量のディスク記憶デバイ

ス検索するためにデバイス番号1（1010）の記憶容量を変数xに格納する（ステップ3050）。変数xに格納された値と次のディスク記憶デバイスの記憶容量とを比較して（ステップ3060）、比較した結果、次のディスク記憶デバイスの記憶容量の方が小さい場合は次のディスク記憶デバイスの記憶容量を変数xに格納する（ステップ3070）。変数xに格納されていた値の方が多い場合は、ステップ3060で比較した次のディスク記憶デバイスの、その次にも比較すべきディスク記憶デバイスが存在するかを判定する（ステップ3080）。これは、コピーした装置容量テーブルにステップ3060で比較した次のディスク記憶デバイスの、その次のディスク記憶デバイス記憶容量が“0”になっていないかをチェックすることによる。もしも、記憶容量が“0”になっている場合には、つまり次に比較すべきディスク記憶デバイス存在しない場合には、エリア番号1をエリア番号格納場所2090に格納する（ステップ3090）。

【0017】次に、エリア(1)での構成ディスク記憶デバイス数を求めるために、変数yにエリア(1)の構成ディスク記憶デバイス数の初期値1を格納する（ステップ3100）。コピーした装置容量テーブルで記憶容量が“0”でない装置全てを対象とし、前記の最小記憶容量検索処理で求められた最小容量の値x分を減算した記憶容量をコピーした装置容量テーブルの各ディスク記憶デバイス対応に書込む（ステップ3110）。言い換えれば、ここで、エリア分割済み記憶容量を削減する処理をする。そこで、コピーした装置容量テーブルには、記憶容量“0”のディスク記憶デバイスと残りの記憶容量が“0”でないディスク記憶デバイスが登録されている。このステップ3110で削減した状態において、次のディスク記憶デバイスの記憶容量が“0”になっていないかをチェックする（ステップ3130）。次のディスク記憶デバイスがある場合はエリア(1)の構成ディスク記憶デバイス数が格納されている変数yを1カウントアップする（ステップ3120）。そして、次にディスク記憶デバイスがなくなるまでこれを続けて、ディスク記憶デバイスがなくなった場合、変数yの値をエリアテーブル2060内のディスク記憶デバイス数格納場所2110に格納する（ステップ3140）。

【0018】次に、エリア(1)で使用している各ディスク記憶デバイスの使用空間情報、つまり、各ディスク記憶デバイスの先頭ブロックアドレス（物理アドレス）と最終ブロックアドレス（物理アドレス）をエリア情報テーブル2060において、それぞれのディスク記憶デバイスのデバイス番号に対応させて欄2130，欄2140に格納する処理を行なう。まず、先頭アドレスは合計使用記憶容量に対応して決定される値zを格納する（ステップ3150）が、このときの使用容量は“0”になっているので、“0”が記録されることになる。エリア(1)は、最初に構築されるものであり、また、合計使用記憶容量zは初期化されているので、どのディスク記憶デバイスも先頭ブロックアドレスは“0”である。

【0019】各ディスク記憶デバイスの使用最終アドレスは、各ディスク記憶デバイスの先頭ブロックアドレスに最小記憶容量検索処理で求められた最小容量値を加算した値、つまり、（先頭ブロックアドレス値＝z）＋変数xとなるので、その値を格納する（ステップ3160）。変数yに格納されているエリア(1)での使用ディスク記憶デバイス数分、すなわち、使用ディスク記憶デバイス数yから1を減算した値が0以上（ステップ3170）である間は、前記のステップ3150からステップ3170をループして、エリアテーブルの装置別使用空間格納場所2140～2146に順次先頭アドレスと最終アドレスとを格納していく（ステップ3180）。

【0020】次に、合計使用記憶容量zにz＋変数xを代入して、これを次の先頭アドレスzとして記憶する（ステップ3180）。ここまでの処理でエリア(1)を構築するためのエリア情報テーブル2060を作成することができる。次に、コピーした装置容量テーブルのディスク記憶デバイス記憶容量が“0”になっているディスク記憶デバイスの数が1か否かを判定し（ステップ3190）、2以上であるときには、ステップ3050へと戻り、次のエリア分割処理を同様な処理を繰り返すことで行う。これをディスク記憶デバイスの数が1になるまでくり返し行ない、繰り返しの都度エリア番号をインクリメントすることでエリア情報テーブル2060が完成し、エリア分割が終了する。これによりエリア(1)とエリア(2)とによる2つのアレイ記憶領域が形成される。なお、ディスク記憶デバイス記憶容量が“0”になっているディスク記憶デバイスの数が1であるときには、冗長データが記憶できないので、このエリアは使用されない。

【0021】次に、複数の記憶装置を物理的に並列接続させた状態の記憶装置を一つの仮想記憶装置として論理的に認識し、アレイ型記憶装置として動作させる仮想記憶装置方式について説明する。これは、記憶容量の小さい記憶デバイスが複数存在するような場合のエリア設定について大きなアドレス空間のエリアを設定することができる点で有効な処理になる。まず、図3を用いて仮想記憶装置について説明する。7000は、物理的に並列接続されたディスク記憶デバイス装置を表わしている。7010は、新規に構築される1個の論理ディスク記憶デバイス（ロジカルユニット）を表わしている。7020は、図7に対応する増設後のディスクアレイ装置の未使用領域を表わしている。7030は、論理的に直列接続された仮想記憶デバイスを表わしている。

【0022】通常アレイ型記憶装置において記憶容量増加の目的で増設する記憶デバイスは、図3の上に示すように、アレイ型記憶装置7000は、実ディスク記憶デ

バイス（以下実デバイスという。）１０１０～１０１３（ＲＤ1～ＲＤ4）に対して新たに増設した実デバイス４０８０（ＲＤ5）とがドライブＩ／Ｆ１００３に物理的に並列接続される。しかし、この状態では、未使用領域は、増設前と変わらず無駄な領域ができてしまう。そこで、図３の下に示すように物理的に接続された状態から増設分実デバイス４０８０（ＲＤ5）と実デバイス１０１３（ＲＤ4）とを論理的に直列接続し、さらに論理的に直列接続された実デバイス１０１３（ＲＤ4）と実デバイス４０８０（ＲＤ5）を１つの仮想記憶デバイスとして認識し、論理的に連続した一つの記憶領域（１個のディスク記憶デバイス）として扱う。このデバイスが論理ディスク記憶デバイスとしての仮想デバイス７０３０である。　なお、以上のような仮想デバイスの設定と管理は、仮想デバイス設定管理プログラム２０３６により行われる。

【００２３】図４は、エリア分割方式とこの仮想記憶装置方式を併用した時に、新たに制御用メモリ２０４０内に作成される仮想記憶装置管理テーブル４０００と、装置容量テーブル２０５０、そしてエリア情報テーブル４０１０とを示している。これらテーブルは、ディスク記憶デバイスの分割状態及び仮想ディスク記憶デバイス（以下仮想デバイスという。）として認識されている状態を表わしている。仮想記憶装置管理テーブル４０００は、実デバイス（ＲＤ1～ＲＤ5）の記憶容量情報が格納されている装置容量テーブル２０５０を元に仮想デバイス設定管理プログラム２０３６により作成される。これには仮想デバイスの詳細情報が格納される。エリア情報テーブル４０１０は、図２のエリア情報テーブル２０６０のデバイス数の欄２１１０が仮想デバイス個数の欄２１５０になり、デバイス番号欄が仮想デバイス番号（ＶＤ）欄２１６０になり、さらに、ＲＤ（実デバイス）の個数欄２１７０、ＲＤ番号欄５１８０が加えられ、仮想デバイス（ＶＤ）と実デバイス（ＲＤ）とを管理するテーブルである。

【００２４】ここで、エリア４０２０は、仮想デバイスを含み、エリア分割されたエリア(1)の領域を示す。エリア４０３０は、仮想デバイス（ＶＤ）を含み、エリア分割されたエリア(2)の領域を示す。４０４０～４０７０は、それぞれ仮想デバイス（ＶＤ）である。また、４０８０は、ディスクアレイ記憶装置の記憶容量を増加する目的で新たに増設された実デバイス（ＲＤ）である。まず始めに、装置容量読込みプログラム２０３２により、接続されている実デバイス（ＲＤ）の記憶容量を全て読出して、装置容量テーブル２０５０を制御用メモリ２０３０に作成する。次に装置容量テーブル２０５０と、オペレータによって指定される仮想デバイスの個数、この実施例では、仮想デバイスＶＤ１，ＶＤ２，ＶＤ３，ＶＤ４の４個と、オペレータによって指定されるそれぞれの仮想デバイスに割当てる実デバイスの情報、この実施例では仮想デバイスＶＤ４が実デバイス１０１３，４０８０（ＲＤ4，ＲＤ5）から構成され、そのほかは、それぞれの実デバイスから構成される。これらの情報を入力して、この入力データに応じて仮想デバイス設定管理プログラム２０３６により仮想記憶装置を管理するエリア情報テーブル４０１０が制御用メモリ２０４０内に作成される。次に、仮想デバイス設定管理プログラム２０３６によりエリア分割プログラム２０３１がコールされて、１仮想デバイス（ＶＤ）を１ディスク記憶デバイスとして扱い、前記実施例のエリア情報テーブル作成手順に従って、１つのエリア内における仮想デバイス単体の記憶容量と、仮想デバイス数と、１つのエリア内の１つの仮想デバイスにおける実デバイスの使用領域情報を求めてエリア情報テーブル４０１０を制御用メモリ２０４０内に作成する。このときには、仮想デバイス数の欄２１５０、仮想デバイス番号（ＶＤ）欄２１６０、ＲＤ（実デバイス）の個数欄２１７０、ＲＤ番号欄５１８０は、それぞれ入力情報に応じてエリア分割プログラム２０３１により生成され、それぞれの情報が記録される。

【００２５】次に、こうして構築したエリア分割方式と仮想記憶装置方式を併用した場合のアレイ型の記憶デバイスを例として上位処理装置１００１からのアクセス要求に対する書込み処理及び読出し処理の動作について図５に従って説明する。図５は、上位処理装置１００１からの書込みアクセス要求に対して実デバイス番号と確定した実デバイス内に格納されているブロックアドレス（物理アドレス）を決定する処理動作を示したものである。５０００は、入出力制御プログラム２０３５による処理装置１００１からのアクセス要求受信を示す。その内容の一例として書込み要求を示している。５０１０は、処理装置１００１から見た、この実施例におけるディスクアレイ装置の論理アドレス空間を示すものである。

【００２６】位置決定プログラム２０３４は、処理装置１００１からのアクセス要求を受入れて、指定されたブロックアドレスに対して所定の演算式に従って演算処理をして記憶デバイスを決定してその物理的な位置（アドレス）を算出する。入出力制御プログラム２０３５は、その位置においてデータの読み／書きを行うとともに、処理装置１００１のデータ書込み要求に対しては処理装置１００１の書込みデータに対して冗長データを生成する処理と、位置決定プログラム２０３４により決定された位置を基準として、書込みデータと冗長データとをディスク記憶デバイスに分散して書込む処理、処理装置１００１からのデータ読出し要求に対しては必要なデータを位置決定プログラム２０３４により決定された位置を基準としてディスク記憶デバイスから読出して、処理装置１００１に転送する処理、障害などによりディスクデバイス上のデータに直接アクセスできないような異常に

おいては、冗長データを用いてアクセスできないデータを復元する処理を行うが、冗長データの生成や冗長データによるデータの復元の処理は、アレイ型記憶装置において、従来から行われている処理であるので、その詳細な説明をここでは省略する。また、ディスク記憶デバイスに分散して書込む処理とは、データをストライピングサイズごとに記憶デバイスを更新しながら順次書込む処理である。したがって、所定のサイズのデータは、ストライピングサイズで各記憶デバイスごとに分散して配置される。そして、その次の記憶デバイスに上長データが書込まれる。なお、冗長データは、データ記憶デバイスとは異なる、データの書込みの都度決定される記憶デバイスに記憶されるものであるので、これを記憶する記憶デバイスがデータを記憶するデバイスのほかに必要になる。そこで、アレイ型記憶装置のデータ記憶容量は、実際にエリア管理されるディスク記憶デバイスよりも１つ少ないデバイスがデータ記憶デバイスになる。また、ここでは、冗長データのシンドロームビットをパリティビットとしてこれを記憶するデバイスをパリティ記憶デバイス（パリティＤＶ）とし、データ記憶デバイスをデータＤＶとして説明する。

【００２７】位置決定プログラム２０３４は、次の計算式５２４０に従ってアクセス要求ブロックアドレスのパリティデータが存在する仮想デバイスの番号ＰＮを算出する。

ＰＮ＝ＶＤ－（（（ＬＢＡ／Ｓ）／Ｄ）％ＶＤ）

ただし、Ｓは、ストライピングサイズ（符号５２００）、ＶＤは、エリア単位の構成仮想デバイス（ＶＤ）の数（符号５２１０）、Ｄは、データを記憶するエリア単位の構成データＶＤ数（符号５２２０）であって、Ｄ＝ＶＤ－１による。このＤ＝ＶＤ－１により冗長データを除く、実データの記録デバイス数が算出される。ＬＢＡは、論理ブロックアドレス値（符号５２３０）である。また、％は除算した結果の余りを算出する演算子である。ＬＢＡ／Ｓにより、論理アドレスについて物理記録単位で換算したカウント総数が算出され、（ＬＢＡ／Ｓ）／Ｄにより実際にデータを書込し／読出しする、あるエリア単位でのデバイスの物理アドレス数が算出される。（（ＬＢＡ／Ｓ）／Ｄ）％ＶＤ）により実際にデータを書込／読出する最後のデバイス番号が算出される。その結果として、アクセス要求ブロックアドレスのパリティデータが存在する仮想デバイスが何番目であるかが、前記式により番号ＰＮとして算出される。

【００２８】また、位置決定プログラム２０３４は、次の計算式５２５０に従ってデータを書込／読出するデバイス番号（データＶＤ）の情報ＤＤを算出する。

ＤＤ＝（ＬＢＡ／Ｓ）％Ｄ＋１

（ＬＢＡ／Ｓ）％Ｄによりあるエリア単位においてデータを書込／読出するデバイス最後のデバイス番号が決定され、アクセス要求ブロックアドレスのデータが存在する仮想デバイスが何番目であるかが、前記式により情報ＤＤとして算出される。さらに、位置決定プログラム２０３４は、次の条件付き計算式５２６０に従ってデータを書込／読出するデバイス（データＶＤ）のデバイス番号ＤＮを算出する。

ＤＮ＝ｉｆ（ＤＤ＞＝ＰＮ）

ＤＮ＝ＤＤ＋１

ｅｌｓｅ

ＤＮ＝ＤＤ

これにより、パリティデータが存在する仮想デバイスを含めて仮想デバイスの番号を算出する。そして、位置決定プログラム２０３４は、次の計算式５２７０に従ってアクセスブロックアドレスＡＢＡを算出する。

ＡＢＡ＝（（（ＬＢＡ／Ｓ）／Ｄ）×Ｓ）＋（ＬＢＡ％（ＬＢＡ／Ｓ））

この式により実デバイスの物理アドレスが算出される。

【００２９】次に、処理装置１００１から所定のブロックアドレスへの書込みを例として全体的な動作を説明する。入出力制御プログラム２０３５は、ブロックアドレス０ｘ１０００００番地に書込みアクセス要求を処理装置１００１から受信（５０００）した場合、エリア情報発生プログラム２０３３をコールし、このプログラムにより、始めに、エリアテーブルの最大エリアサイズと要求を受けたブロックアドレスからエリア番号を算出する。このプログラムは、エリア(1)の最大エリアサイズ０ｘ１２０３１００番地（５１３０）よりアクセス要求ブロックアドレス０ｘ１０００００番地の方が小さいことから、要求ブロックアドレスはエリア(1)内（５１２０）に存在すると特定する。入出力制御プログラム２０３５は、次にブ位置決定プログラム２０３４をコールして、このプログラムによりストライピングサイズ：Ｓ（５２００）にはエリア情報テーブル４０１０内のストライピングサイズ５１２バイト（５１１０）を代入する。そして、エリア単位の構成仮想デバイス数：ＶＤ（５２１０）には、エリア(1)の構成仮想デバイス数４（５１４０）を代入し、エリア単位の構成データ仮想デバイス数：Ｄ（５２２０）には、ＶＤ－１＝３を代入する。さらに、アクセス要求ブロックアドレスＬＢＡ（５２３０）には、アクセス要求ブロックアドレスの０ｘ１０００００番地を代入する。

【００３０】これらの値をアクセス要求ブロックアドレスが存在するデータのパリティ仮想デバイス番号を特定するための計算式（５２４０）とアクセス要求ブロックアドレスが存在するデータのディスク記憶デバイスが何番目かを特定する計算式５２５０に代入してそれぞれの値を算出する。前記計算式により算出されたデータディスク記憶デバイス値：ＤＤは、パリティ仮想デバイス位置によって仮想デバイス位置が変わるため、計算式（５２６０）によりデータディスク記憶デバイス番号を特定する。前記の実施例のディスクアレイ装置では、ＰＮ＝

２，Ｄ＝３，ＤＮ＝４となる。

【００３１】次に図６に従って、入出力制御プログラム２０３５による、前記の算出された情報を元にしてアクセスを行なうまでの手順を説明する。始めに、仮想デバイス内の何番目の実デバイスであるかを記憶する変数ＲＤｃｏｕｎｔを初期する（ステップ６０００）。次に、エリアは(1)と判明していることと、前記で算出した情報ＤＮが４であることから、エリア情報テーブル４０１０のエリア(1)、仮想デバイス番号４、ＲＤｃｏｕｎｔ：１の先頭アドレスを変数ｘに（ステップ６０１０）入れ、最終アドレスを変数ｙに格納する（ステップ６０２０）。次に、アクセスアドレスを計算式（５２７０）で算出する。前記の実施例では、０ｘ５５５４００番地という値が算出される（ステップ６０３０）。算出されたアクセスブロックアドレスＡＢＡが最大エリアサイズを越えている可能性があるので、一つ前のエリアの最大エリアサイズを減算する（ステップ６０４０）。

【００３２】エリア(1)へのアクセスであるため“０”をここでは減算する。次に上記で算出したアクセスブロックアドレス（０ｘ５５５４００番地）に先頭ブロックアドレスｘを加算した値が最終ブロックアドレスｙを越えているか、つまり、アクセスブロックアドレスの値が使用している実デバイス容量を越えていないかチェックする（ステップ６０５０）。もし越えている場合は、実デバイス先頭アドレス：ｘにアクセスブロックアドレス：ＡＢＡを加算した値からその実デバイスの最終アドレス：ｙを減算した値をＡＢＡに格納し（ステップ６０６０）、ＲＤｃｏｕｎｔを１加算する（ステップ６０７０）処理をアクセス領域を発見するまでくり返す。越えていない場合、つまりアクセス領域を発見した場合は、エリア情報テーブル４０１０とＲＤｃｏｕｎｔから実デバイス番号を求め、変数ＲＤＮｏに格納する（ステップ６０８０）。実デバイスへのアクセスは、求められた実デバイスの先頭アドレスｘに求められたアクセスブロックアドレスＡＢＡを加算した値のアドレスに行なう。

【００３３】このようにして得た情報から前記の実施例の場合は、実デバイス番号（ＲＤＮｏ：５）のアクセスブロックアドレス（ＡＢＡ：０ｘ１５５４００番地）にアクセスすることで処理装置１００１からの要求に正しく応答することが可能である（ステップ６０９０）。なお、処理装置１００１からの読込みアクセス要求に対しても同様の処理を行なうことにより正しいアドレスへのアクセス処理が可能である。以上、この発明を実施例に基づき具体的に説明したが、この発明は、前記実施例に限定されるものではなく、その要旨を逸脱しない範囲で種々変更しえることは言うまでもない。

【００３４】この実施例によれば、このように、アレイ型記憶装置において、制御装置の配下に接続された各記憶デバイスの記憶容量を読出し、異なる記憶容量の記憶デバイスにおいても、エリア分割を用いてエリア分割情報作成し、さらには、エリア分割方式と記憶デバイスを論理的に直列に接続する仮想記憶デバイスとの併用により、エリア分割情報内に仮想記憶デバイス情報を盛り込み、処理装置からのアクセス要求に対しては、計算式によって実記憶デバイス番号、アクセスアドレスの算出を容易に行なえるので、アレイ型記憶装置において、異なる記憶容量を持つ記憶デバイスを有効に活用することができる。

【００３５】

【発明の効果】この発明にあっては、アレイ形式の記憶装置システムにおいて、搭載する記憶デバイス単体の記憶容量が同一であることを考慮してシステムを構築するという制限をなくし、異なる記憶容量を持つ記憶デバイスを最大限に利用することが可能になる。また、記憶容量増加の目的で増設した増設記憶デバイスの記憶容量が増設以前に搭載されている記憶デバイスの記憶容量と同一でなければ最大限、記憶容量を有効に利用できない制限をなくし、増設する記憶デバイスの記憶容量が増設以前に搭載されている記憶デバイスの記憶容量と異なっていても記憶デバイスを最大限に活用できる。

【図面の簡単な説明】

【図１】図１は、この発明のアレイ型記憶装置を適用した一実施例のディクアレイ記憶システムの機能ブロック図である。

【図２】図２は、図１の実施例におけるエリア情報テーブルの生成処理のフローチャートである。

【図３】図３は、図１の実施例において、物理的に並列接続された記憶デバイスを論理的に直列接続した時の仮想記憶デバイスの認識状態の説明図である。

【図４】図４は、異なる記憶容量を持つ記憶装置を論理的に直列に接続した場合の図１の実施例における、新たに制御用メモリ内に作成するエリア情報テーブルとディスク記憶デバイスの分割状態及び仮想デバイスとして認識されている状態の説明図である。

【図５】図５は、上位処理装置からの書込みアクセス要求に対して実記憶デバイス番号と確定した実記憶デバイス内に格納されているブロックアドレスを決定する説明図である。

【図６】図６は、図１の実施例において、計算式により算出された情報を元に実記憶デバイスに対してのアクセス処理を行なうまでの処理のフローチャートである。

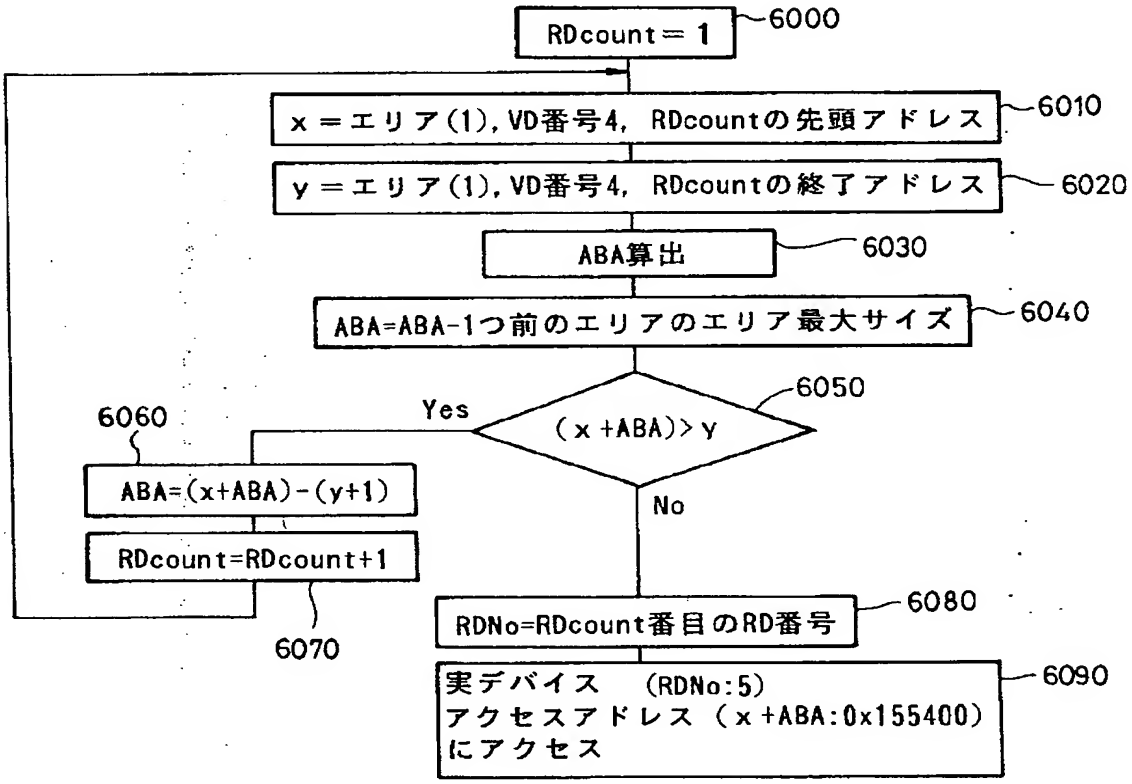
【図７】図７は、一般的なディスクアレイ型記憶システムの全体的な構成図である。

【符号の説明】

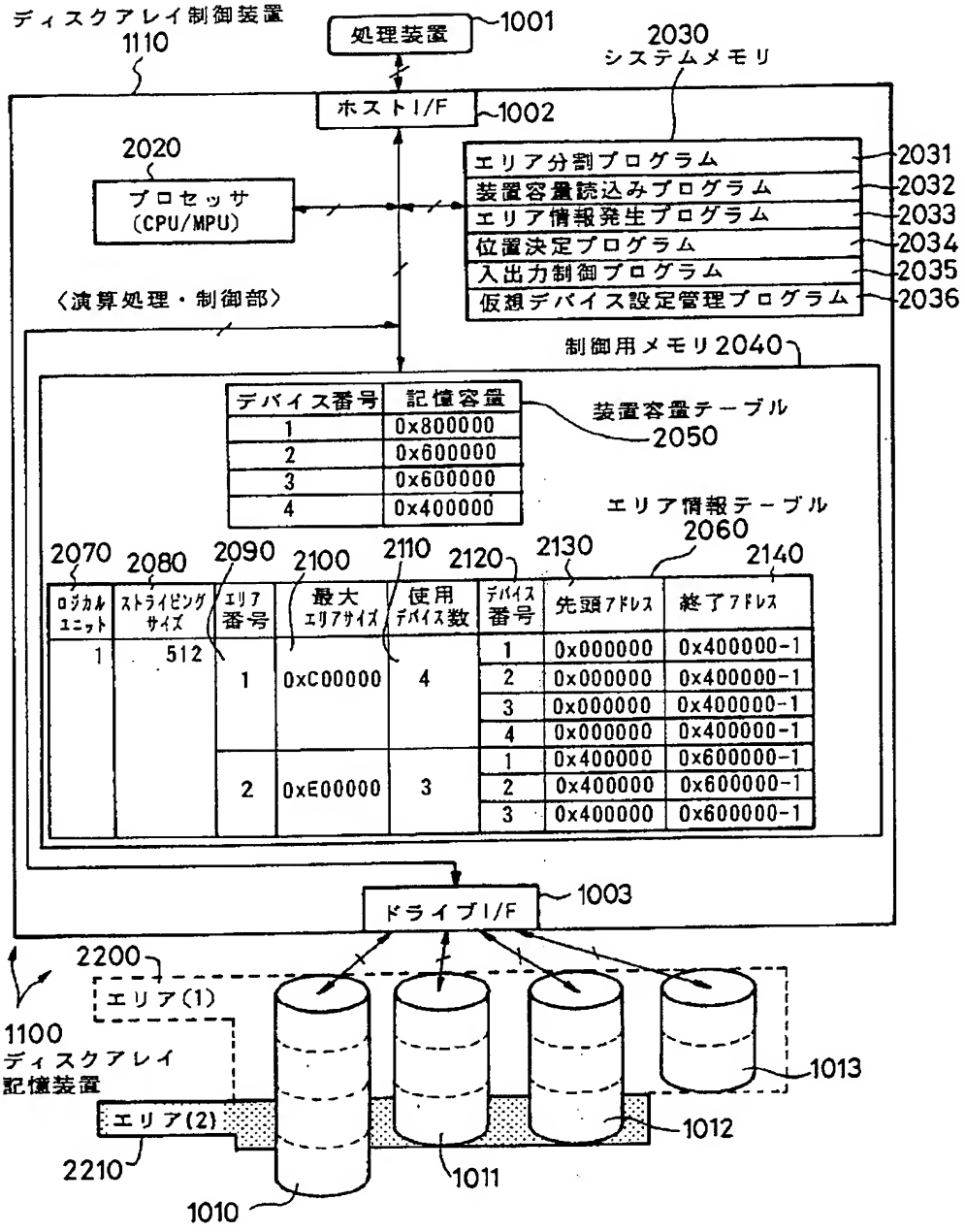
１０００…ディスクアレイ制御装置、１００１…処理装置、１００２ホストＩ／Ｆ、１００３…ドライブＩ／Ｆ、１００４，２０３０…制御用メモリ、１００５…プロセッサ（ＣＰＵ／ＭＰＵ）、１０１０～１０１３…ディスク記憶デバイス、１０２０…ロジカルユニット、１０２１…未使用領域、１０３０…ディスクアレイ装置、

２０３１…エリア分割プログラム、２０３１…装置容量読込みプログラム、２０３３…エリア識別プログラム、２０３４…位置決定プログラム、２０５０…装置容量テーブル、２０６０…エリア分割情報テーブル、２０７０…ロジカルユニット番号格納場所、２０８０…ストライピングサイズ格納場所、２０９０…エリア番号格納場所、２１００…最大エリアサイズ格納場所、２１１０…使用ディスク記憶デバイス数格納場所、２１２０～２１８０…ディスク記憶デバイス使用領域格納場所、２２００，２２１０…分割したエリア、４０００…仮想記憶装置管理テーブル、４０１０…仮想記憶装置を考慮したエリア情報テーブル、４０２０，４０３０…仮想記憶装置を考慮し分割したエリア、４０４０～４０７０…仮想デバイス、４０８０…増設したディスク記憶デバイス、５０１０…処理装置から見た論理空間、５２４０…パリティディスク記憶デバイス番号算出計算式、５２５０…データディスク記憶デバイス算出計算式、５２６０…データディスク記憶デバイス番号算出計算式、５２７０…アクセスブロックアドレス算出計算式、７０３０…仮想記憶装置。

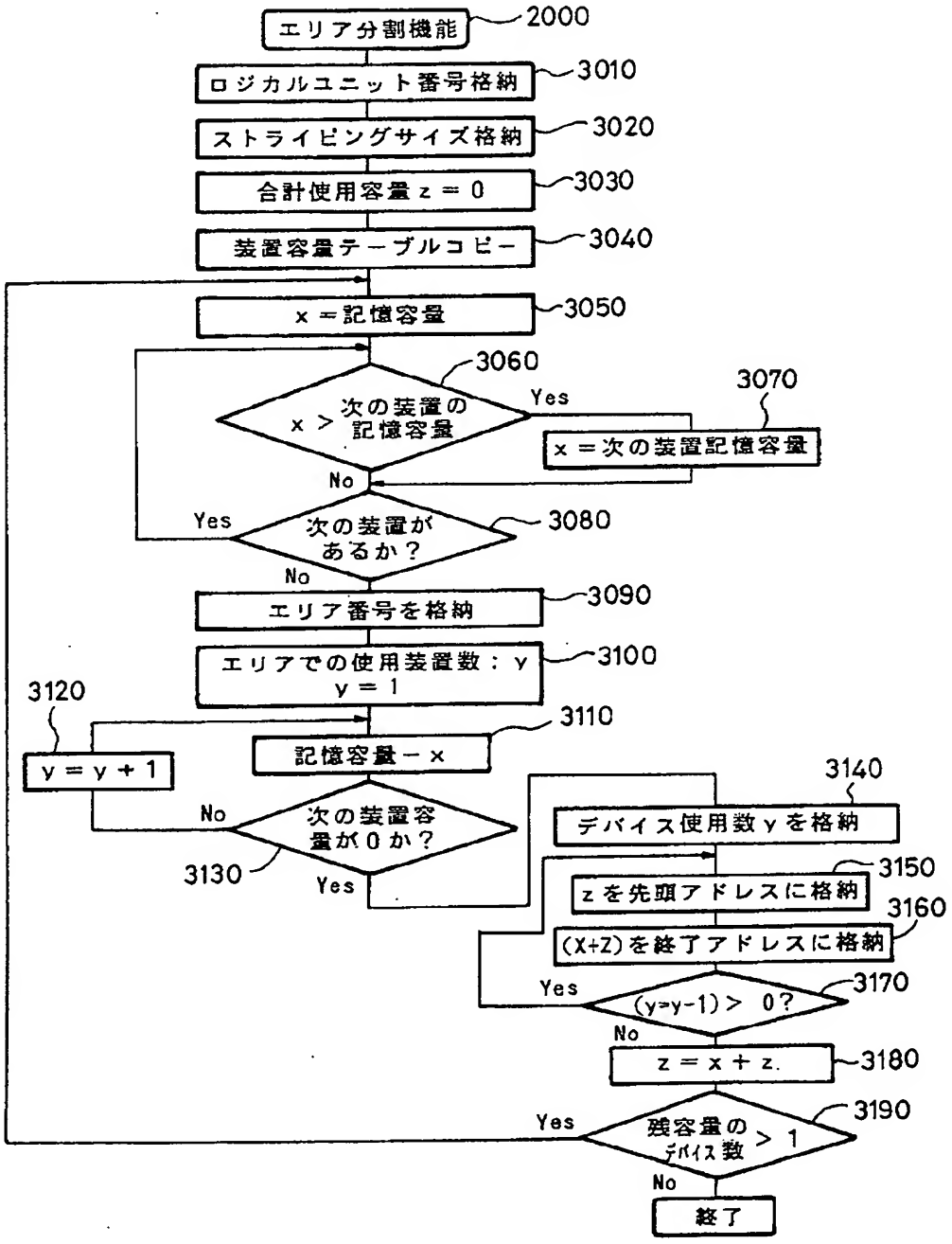
【図６】



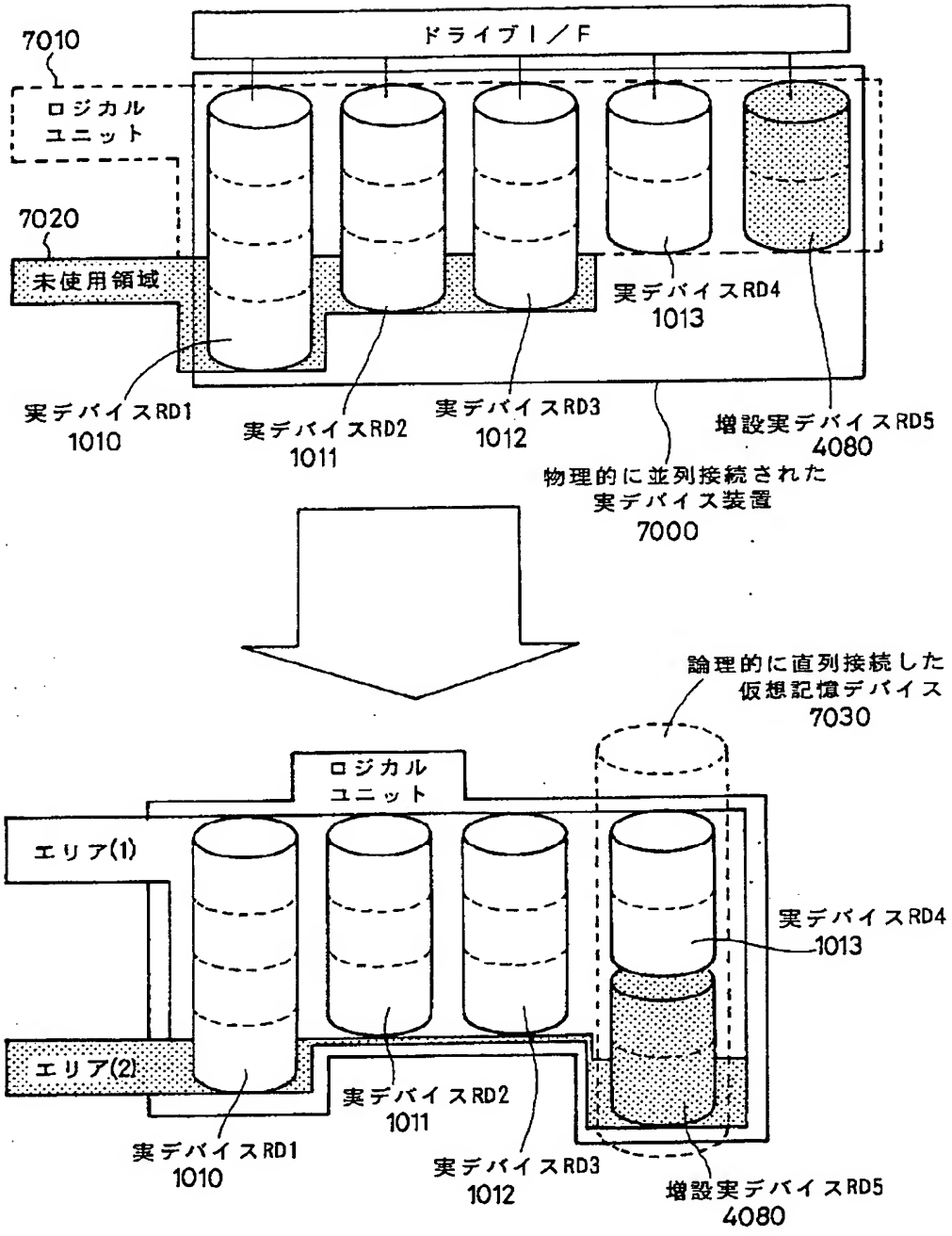
【図1】


ディスクアレイ制御装置
1110
処理装置
1001
2030
システムメモリ
ホストI/F
1002
2020
プロセッサ
(CPU/MPU)
エリア分割プログラム
装置容量読込みプログラム
エリア情報発生プログラム
位置決定プログラム
入出力制御プログラム
仮想デバイス設定管理プログラム
2031
2032
2033
2034
2035
2036
〈演算処理・制御部〉
制御用メモリ2040
デバイス番号
記憶容量
1
0x800000
2
0x600000
3
0x600000
4
0x400000
装置容量テーブル
2050
エリア情報テーブル
2060
2070
2080
2090
2100
2110
2120
2130
2140
ロジカルユニット
ストライピングサイズ
エリア番号
最大エリアサイズ
使用デバイス数
デバイス番号
先頭アドレス
終了アドレス
1
512
1
0xC00000
4
1
0x000000
0x400000-1
2
0x000000
0x400000-1
3
0x000000
0x400000-1
4
0x000000
0x400000-1
2
0xE00000
3
1
0x400000
0x600000-1
2
0x400000
0x600000-1
3
0x400000
0x600000-1
ドライブI/F
1003
2200
エリア(1)
1100
ディスクアレイ
記憶装置
エリア(2)
2210
1010
1011
1012
1013

【図2】


エリア分割機能
2000
ロジカルユニット番号格納
3010
ストライピングサイズ格納
3020
合計使用容量 z = 0
3030
装置容量テーブルコピー
3040
x = 記憶容量
3050
3060
x > 次の装置の記憶容量
Yes
3070
x = 次の装置記憶容量
No
Yes
次の装置があるか？
3080
No
エリア番号を格納
3090
エリアでの使用装置数：y
y = 1
3100
3120
y = y + 1
記憶容量－x
3110
No
次の装置容量が0か？
3130
Yes
3140
デバイス使用数 y を格納
z を先頭アドレスに格納
3150
(X+Z) を終了アドレスに格納
3160
Yes
(y=y-1) > 0？
3170
No
z = x + z.
3180
Yes
残容量のデバイス数 > 1
3190
No
終了

【図３】


7010
ドライブＩ／Ｆ
ロジカル
ユニット
7020
未使用領域
実デバイスRD4
1013
実デバイスRD1
1010
実デバイスRD2
1011
実デバイスRD3
1012
増設実デバイスRD5
4080
物理的に並列接続された
実デバイス装置
7000
論理的に直列接続した
仮想記憶デバイス
7030
ロジカル
ユニット
エリア(1)
エリア(2)
実デバイスRD4
1013
実デバイスRD2
1011
実デバイスRD1
1010
実デバイスRD3
1012
増設実デバイスRD5
4080

【図4】

制御用メモリ 2040

仮想記憶装置管理テーブル 4000

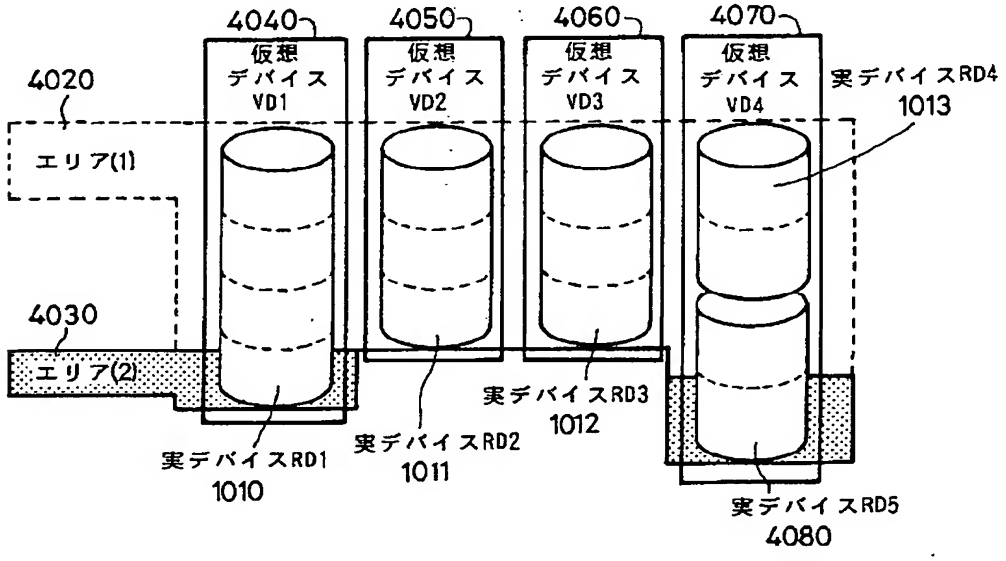
VD番号	VD容量	RD個数	RD番号	容量
1	0x800000	1	1	0x800000
2	0x600000	1	2	0x600000
3	0x600000	1	3	0x600000
4	0x800000	2	4	0x400000
			5	0x400000

装置容量テーブル 2050

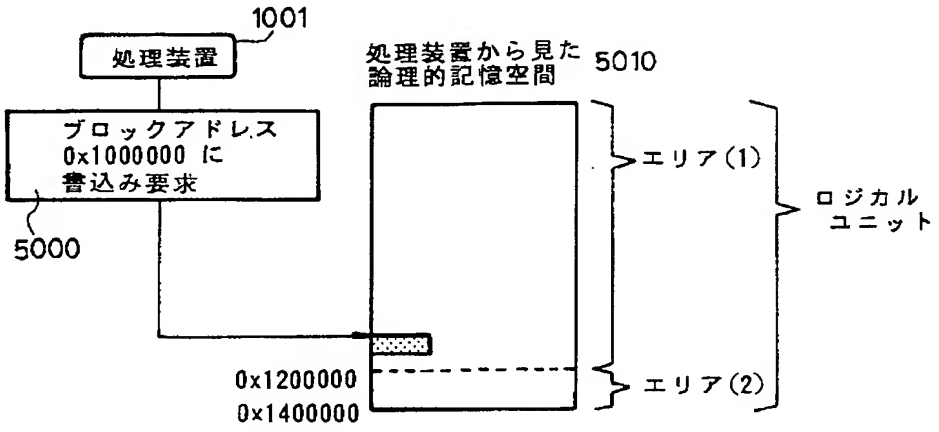
デバイス番号	記憶容量
1	0x800000
2	0x600000
3	0x600000
4	0x400000

エリア情報テーブル 4010

ロジカルユニット (2070)	ストライピングサイズ (2080)	エリア番号 (2090)	最大エリアサイズ (2100)	仮想デバイス(VD)情報					
				VD個数 (2150)	VD番号 (2160)	実デバイス(RD)情報			
						RD個数 (2170)	RD番号 (2180)	先頭アドレス	終了アドレス (2140)
1	512	1	0x1200000	4	1	1	1	0x000000	0x600000-1
					2	1	2	0x000000	0x600000-1
					3	1	3	0x000000	0x600000-1
					4	2	4	0x000000	0x400000-1
							5	0x000000	0x200000-1
		2	0x1400000	2	1	1	1	0x600000	0x800000-1
					4	1	5	0x200000	0x400000-1



【図5】



ストライピングサイズ : S ——— 5200
エリア単位の構成VD数 : VD ——— 5210
エリア単位の構成データVD数 : D=VD-1 ——— 5220
論理ブロックアドレス : LBA ——— 5230

パリティVD番号 : PN=VD-(((LBA/S)/D)%VD) ——— 5240
データVD : DD=(LBA/S)%D+1 ——— 5250
データVD番号 : DN=if(DD>=PN)
DN=DD+1
else
DN=DD ——— 5260
ｱｸｾｽﾌﾞﾛｯｸｱﾄﾞﾚｽ : ABA=(((LBA/S)/D)×S)+(LBA%(LBA/S)) ——— 5270

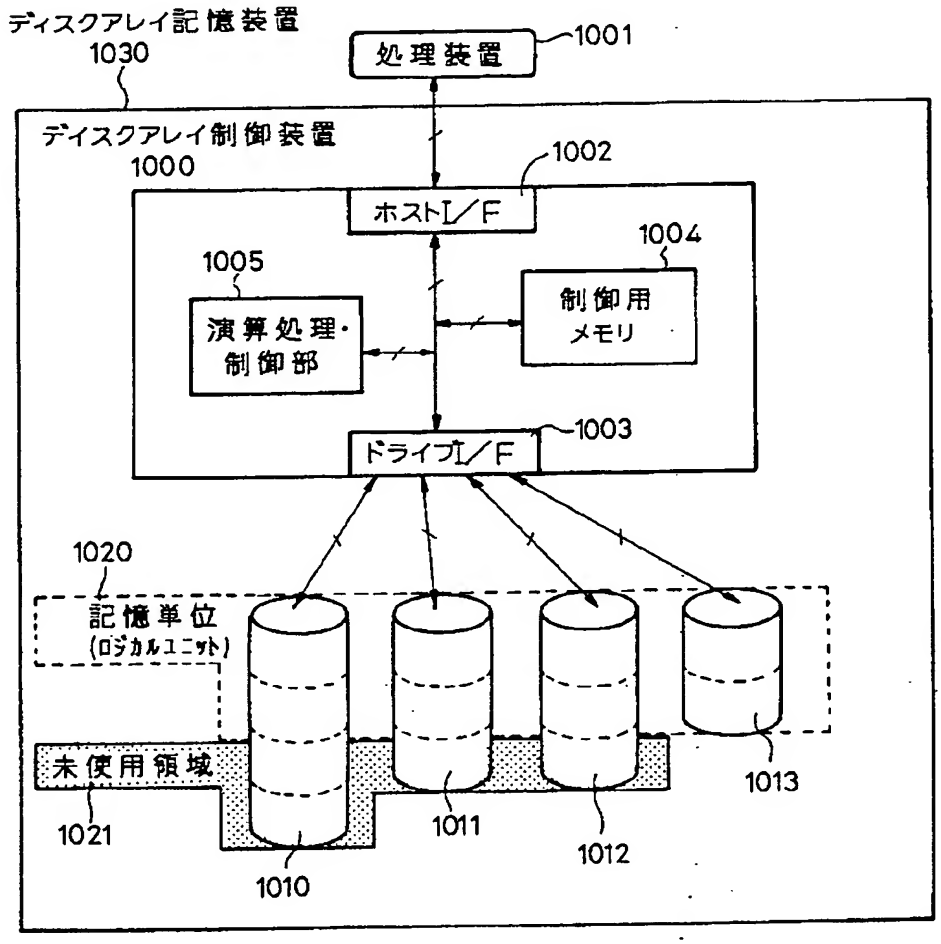
注） A/B:Aを Bで除算した商
A%B:A を Bで除算した余り

エリア情報テーブル 4010

ﾛｼﾞｶﾙ ﾕﾆｯﾄ	ｽﾄﾗｲﾋﾟﾝｸﾞ ｻｲｽﾞ	ｴﾘｱ 番号	最大 ｴﾘｱｻｲｽﾞ	仮想デバイス(VD)情報					
				VD 個数	VD 番号	実ﾃﾞﾊﾞｲｽ(RD)情報			
						RD 個数	RD 番号	先頭ｱﾄﾞﾚｽ	終了ｱﾄﾞﾚｽ
1	512	1	0x1200000	4	1	1	1	0x000000	0x600000-1
					2	1	2	0x000000	0x600000-1
					3	1	3	0x000000	0x600000-1
					4	2	4	0x000000	0x400000-1
							5	0x000000	0x200000-1
		2	0x1400000	2	1	1	1	0x600000	0x800000-1
					4	1	5	0x200000	0x400000-1

5100 5110 5120 5130 5140

【図7】


ディスクアレイ記憶装置
1030
処理装置
1001
ディスクアレイ制御装置
1000
1002
ホストI/F
1004
1005
演算処理・制御部
制御用メモリ
1003
ドライブI/F
1020
記憶単位
(ロジカルユニット)
未使用領域
1021
1010
1011
1012
1013